

# Characterization of Clinically-Attenuated *Burkholderia mallei* by Whole Genome Sequencing: Candidate Strain for Exclusion from Select Agent Lists

Steven E. Schutzer<sup>1\*</sup>, Linda R. K. Schlater<sup>2</sup>, Catherine M. Ronning<sup>3</sup>, David DeShazer<sup>4</sup>, Benjamin J. Luft<sup>5</sup>, John J. Dunn<sup>6</sup>, Jacques Ravel<sup>7,8</sup>, Claire M. Fraser-Liggett<sup>7,8</sup>, William C. Nierman<sup>3,9</sup>

**1** Department of Medicine, University of Medicine and Dentistry - New Jersey Medical School, Newark, New Jersey, United States of America, **2** United States Department of Agriculture, Ames, Iowa, United States of America, **3** J. Craig Venter Institute, Rockville, Maryland, United States of America, **4** U.S. Army Medical Research Institute of Infectious Diseases, Fort Detrick, Maryland, United States of America, **5** Department of Medicine, State University of New York, Stony Brook, New York, United States of America, **6** Biology Department, Brookhaven National Laboratory, Upton, New York, United States of America, **7** Institute for Genome Sciences, Department of Medicine, University of Maryland School of Medicine, Baltimore, Maryland, United States of America, **8** Institute for Genome Sciences, Department of Microbiology and Immunology, University of Maryland School of Medicine, Baltimore, Maryland, United States of America, **9** Department of Biochemistry and Molecular Biology, The George Washington University School of Medicine, Washington, D. C., United States of America

## Abstract

**Background:** *Burkholderia mallei* is an understudied biothreat agent responsible for glanders which can be lethal in humans and animals. Research with this pathogen has been hampered in part by constraints of Select Agent regulations for safety reasons. Whole genomic sequencing (WGS) is an apt approach to characterize newly discovered or poorly understood microbial pathogens.

**Methodology/Principal Findings:** We performed WGS on a strain of *B. mallei*, SAVP1, previously pathogenic, that was experimentally infected in 6 equids (4 ponies, 1 mule, 1 donkey), natural hosts, for purposes of producing antibodies. Multiple high inocula were used in some cases. Unexpectedly SAVP1 appeared to be avirulent in the ponies and mule, and attenuated in the donkey, but induced antibodies. We determined the genome sequence of SAVP1 and compared it to a strain that was virulent in horses and a human. In comparison, this phenotypic avirulent SAVP1 strain was missing multiple genes including all the animal type III secretory system (T3SS) complex of genes demonstrated to be essential for virulence in mice and hamster models. The loss of these genes in the SAVP1 strain appears to be the consequence of a multiple gene deletion across insertion sequence (IS) elements in the *B. mallei* genome. Therefore, the strain by itself is unlikely to revert naturally to its virulent phenotype. There were other genes present in one strain and not the other and vice-versa.

**Conclusion/Significance:** The discovery that this strain of *B. mallei* was both avirulent in the natural host ponies, and did not possess T3SS associated genes may be fortuitous to advance biodefense research. The deleted virulence-essential T3SS is not likely to be re-acquired naturally. These findings may provide a basis for exclusion of SAVP1 from the Select Agent regulation or at least discussion of what else would be required for exclusion. This exclusion could accelerate research by investigators not possessing BSL-3 facilities and facilitate the production of reagents such as antibodies without the restraints of Select Agent regulation.

**Citation:** Schutzer SE, Schlater LRK, Ronning CM, DeShazer D, Luft BJ, et al. (2008) Characterization of Clinically-Attenuated *Burkholderia mallei* by Whole Genome Sequencing: Candidate Strain for Exclusion from Select Agent Lists. PLoS ONE 3(4): e2058. doi:10.1371/journal.pone.0002058

**Editor:** Frederick M. Ausubel, Massachusetts General Hospital, United States of America

**Received** February 4, 2008; **Accepted** March 13, 2008; **Published** April 30, 2008

This is an open-access article distributed under the terms of the Creative Commons Public Domain declaration which stipulates that, once placed in the public domain, this work may be freely reproduced, distributed, transmitted, modified, built upon, or otherwise used by anyone for any lawful purpose.

**Funding:** Work was funded in part by several grants from the National Institutes of Health (N01-AI30071, U01 AI056480, AI063757). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: schutzer@umdnj.edu

## Introduction

An initial approach today to a newly discovered pathogen is to perform whole genome sequencing (WGS). This same approach is relevant for investigations of an understudied high-consequence pathogen such as *Burkholderia mallei*, the cause of glanders. This zoonotic bacterium can kill humans and animals [1,2]. It is classified as a Category B biothreat agent and it has been used as a biological warfare agent [2–8]. Its ease of use against civilians and lack of countermeasures has compelled several agencies to rank it

very high on their priority lists of biothreat agents[9,10]. Its ease of use as a biothreat agent is illustrated by the physician who, even in 1915, grew it out of his Washington DC area home, and distributed it for an attack against US horses destined for Europe as critical transportation in World War I[11]. Humans were also subject to infection. We remain highly vulnerable to this agent because there is no rapid diagnostic assay, no distinctive diagnostic signs, and the incubation period is short, a few days. In addition there is no vaccine, no infection-induced immunity, and limited reliable *in vivo* data on antibiotic efficacy. Overall, our general



Report Documentation Page			Form Approved OMB No. 0704-0188	
<p>Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.</p>				
1. REPORT DATE <b>01 JAN 2008</b>	2. REPORT TYPE <b>N/A</b>	3. DATES COVERED <b>-</b>		
<b>4. TITLE AND SUBTITLE</b> <b>Characterization of clinically attenuated <i>Burkholderia mallei</i> by whole-genome sequencing: candidate strain for exclusion from select agent lists</b>			5a. CONTRACT NUMBER	
			5b. GRANT NUMBER	
			5c. PROGRAM ELEMENT NUMBER	
<b>6. AUTHOR(S)</b> <b>Schutzer, SE Schlater, LRK Ronning, CM DeShazer, D Luft, BJ Dunn, JJ Fraser-Ligget, CM Nierman, WC</b>			5d. PROJECT NUMBER	
			5e. TASK NUMBER	
			5f. WORK UNIT NUMBER	
<b>7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)</b> <b>United States Army Medical Research Institute of Infectious Diseases, Fort Detrick, MD</b>			8. PERFORMING ORGANIZATION REPORT NUMBER <b>TR-08-014</b>	
			10. SPONSOR/MONITOR'S ACRONYM(S)	
			11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
			12. DISTRIBUTION/AVAILABILITY STATEMENT <b>Approved for public release, distribution unlimited</b>	
<b>13. SUPPLEMENTARY NOTES</b> The original document contains color images.				
<b>14. ABSTRACT</b> <p><b>Background:</b> <i>Burkholderia mallei</i> is an understudied biothreat agent responsible for glanders which can be lethal in humans and animals. Research with this pathogen has been hampered in part by constraints of Select Agent regulations for safety reasons. Whole genomic sequencing (WGS) is an apt approach to characterize newly discovered or poorly understood microbial pathogens. <b>Methodology/Principal Findings:</b> We performed WGS on a strain of <i>B. mallei</i>, SAVP1, previously pathogenic, that was experimentally infected in 6 equids (4 ponies, 1 mule, 1 donkey), natural hosts, for purposes of producing antibodies. Multiple high inocula were used in some cases. Unexpectedly SAVP1 appeared to be avirulent in the ponies and mule, and attenuated in the donkey, but induced antibodies. We determined the genome sequence of SAVP1 and compared it to a strain that was virulent in horses and a human. In comparison, this phenotypic avirulent SAVP1 strain was missing multiple genes including all the animal type III secretory system (TTSS) complex of genes demonstrated to be essential for virulence in mice and hamster models. The loss of these genes in the SAVP1 strain appears to be the consequence of a multiple gene deletion across insertion sequence (IS) elements in the <i>B. mallei</i> genome. Therefore, the strain by itself is unlikely to revert naturally to its virulent phenotype. <b>Conclusion/Significance:</b> The discovery that this strain of <i>B. mallei</i> was both avirulent in the natural host ponies, and did not possess TTSS associated genes may be fortuitous to advance biodefense research. The deleted virulence-essential TTSS is not likely to be re-acquired naturally. These findings may provide a basis for exclusion of SAVP1 from the Select Agent regulation or at least discussion of what else would be required for exclusion. This exclusion could accelerate research by investigators not possessing BSL-3 facilities and facilitate the production of reagents such as antibodies without the restraints of select agent regulation.</p>				
<b>15. SUBJECT TERMS</b> <i>Burkholderia mallei</i> , glanders, whole-genome sequencing, SAVP strain, no reversion to phenotype, avirulence				
16. SECURITY CLASSIFICATION OF:		17. LIMITATION OF ABSTRACT <b>SAR</b>	18. NUMBER OF PAGES <b>6</b>	19a. NAME OF RESPONSIBLE PERSON
a. REPORT <b>unclassified</b>	b. ABSTRACT <b>unclassified</b>	c. THIS PAGE <b>unclassified</b>		

knowledge of this understudied pathogen and its disease is limited [2,12–17]. Our desire to decrease our vulnerabilities to this pathogen, as part of the national biodefense efforts, is hampered, in part, by constraints of the Select Agent regulation and the need for BSL-3 facilities. These are appropriately in place for safety reasons, however, the availability of a suitable attenuated surrogate strain would be desirable as it could accelerate *B. mallei* related biodefense research in many non-BSL-3 laboratories.

*B. mallei*[18] is a Gram-negative non-motile aerobic bacteria with a genome of approximately 6 Mb organized in two circular chromosomes. In addition to infecting humans, *B. mallei* can cause acute or chronic fatal contagious zoonotic infections in its natural equine host, such as horses, donkeys, and mules, with a very low infectious dose. Two major potential routes of infection for a biologic attack are aerosol and cutaneous contact. Gastrointestinal ingestion is a common mode of natural infection in equines. The incubation period is typically between 3–6 days but may be longer.

An example of a phenotypic highly virulent strain is *B. mallei* strain ATCC 23344 (China 7). It is highly virulent in its natural hosts, equines, in humans, and in mice and hamster models[14,19–21]. This strain of *B. mallei* contains an animal type III secretion system (T3SS) gene complex which is essential for virulence[15,22]. Genome sequencing (WGS) and analysis of this strain identified a number of other putative virulence factors whose function was supported by comparative genome hybridization and expression profiling of the bacterium in hamster liver *in vivo*[22]. Numerous insertion sequence elements that have mediated extensive deletions and rearrangements of the genome relative to the *B. pseudomallei* genome were found. As part of our interest to study mechanisms responsible for virulence we performed WGS and comparative genomic analysis on strains of *B. mallei* in which we had closely linked information from an actual clinical infection in its natural host or a suitable animal model rather than using an archived strain lacking this information. We intended to compare the gene content of phenotypic virulent to avirulent strains to identify candidate virulence genes. One serendipitous event and an unexpected finding led to what may be a fortuitous observation. The serendipitous event occurred when five of six equids were infected, but not made ill, with what was a previously pathogenic strain of *B. mallei*, designated as SAVP1. At the time of inoculation this strain was still believed to be pathogenic. A sixth equid, a donkey, developed clinical signs only after a massive exposure. The strain appeared avirulent even when administered in escalating doses. The other unexpected finding came while performing comparative genomic analysis between SAVP1 and the strain of *B. mallei* that was currently behaving as virulent in humans and horses. This finding is both informational and fortuitous because it may open the door for SAVP1 to be classified as suitably avirulent for exclusion from Select Agent lists[23] thereby expanding our biodefense research efforts.

## Results

### In vivo characterization of phenotypic avirulent strain of *B. mallei* in natural equine hosts

Based upon our premise that relevant pathogenic differences between strains may be revealed by WGS of a bona-fide phenotypic avirulent strain of *B. mallei*, we wanted to select a strain that was previously evaluated in the natural host. This was accomplished by selecting a strain (SAVP1) that had previously caused disease in a mule (in India) and surprisingly did not produce overt disease when inoculated, first orally and subsequently intravenously, into a mule and 4 ponies, even at escalating

doses approaching  $10^9$  colony forming units (CFU) per ml. The prime purpose of the initial experiment, carried out by one of us (LKS) more than twenty years ago, was to develop a diagnostic assay for *B. mallei* as part of the US Department of Agricultures mission to help prevent the disease from entering the United States through horses. It was not meant as a controlled infectivity study, consequently full clinical description and detailed history of the strain are not available. Notes indicate 4 equids (4 ponies, a mule, and a donkey) were exposed by pharyngeal spray twice to increasing doses of *B. mallei*. When none of the equids became ill or seroconverted by 30 days post exposure, they were challenged intravenously with 5 ml of heavy suspension of organisms. The extreme exposure produced clinical signs only in the donkey, a species known to be more susceptible than horses and mules to *B. mallei*. The donkey was euthanized for humane reasons. This donkey received a high intravenous dose, approaching believed to have approached 5 ml of  $10^9$  cfu/ml – thus the symptoms may have been related to endotoxin. Clinical signs in the ponies and the mule were either minimal (fever) or nonexistent. All 6 equids produced antibodies to *B. mallei* as a result of the intravenous exposure. The expected clinical signs of morbid infection were not present in the ponies or the mule. At the time of the necropsy, there was little evidence of systemic pathology. This included the absence of the common sinopulmonary findings of purulent nasal discharge, ulcerations, pleural effusion, pulmonary edema, congestion, or pneumonia. The sole pulmonary involvement was restricted to a few minimal bronchopulmonary granulomatous lesions which did not grow *B. mallei* on culture. From a microbiologic vantage it was not possible to ascertain if current avirulence was an attenuation effect that arose from passage in the mule before it was grown in culture and re-inoculated or if it was an effect of the *in vitro* culture. What is important here is the observation that this particular strain had avirulent/attenuated behavior in this natural host experiment. We performed WGS on DNA extracted from the same or near (low-passage) generation growth of the *B. mallei* used in this experiment.

### Detection of genes present or absent in an avirulent strain from a natural equine host in comparison to a known virulent strain

The SAVP1 strain was subjected to WGS and the resulting data compared those from the virulent ATCC 23344 strain[22] which was known to have caused recent near-death in a human and five horses that had to be euthanized[19,21]. A one way analysis interrogated what genes were present in ATCC 23344 but lacking in the SAVP1 strain. This analysis took all of the ATCC 23344 coding sequences (CDS) and aligned them against the SAVP1 genomic sequence, identifying 582 ORFs that were present in ATCC 23344 but not SAVP1, the vast majority of which exist within a few large gene clusters (See Table 1, Supplementary Table S1, and Figure 1). In contrast, all of the genes shared between the two isolates are at least 99% identical over 99% or more of their length, as determined by blastn.

The most noteworthy difference was the loss of all the animal T3SS associated genes[15] in an apparent IS mediated deletion containing contiguous genes (two sets of T3SS: one consists of 23 genes from BMA\_A1530-A1552 (the Bsa genes) and a second set from BMA\_A1625-A1637 consists of 13 genes, Genbank accession numbers, ATCC 23344 small chromosome- CP000011 and SAVP1- CP000525). These genes are located within the largest of the ATCC 23344-specific gene clusters mentioned above. The absence of animal type T3SS alone could explain its avirulence in the natural host. Differences in the ATCC 23344 and SAVP1 genomes result from multiple inversions and deletions of genome

**Table 1.** Type III secretion proteins encoded by genes present in ATCC 23344 but absent in SAVP1.

Locus	Annotation	5' end	3' end
BMA_A1520	type III secretion chaperone BicP	1650045	1649590
BMA_A1532	type III secretion chaperone BicA	1662375	1661830
BMA_A1533	type III secretion system protein BsaZ	1663706	1662471
BMA_A1534	type III secretion system protein BsaY	1664480	1663710
BMA_A1535	type III secretion system protein BsaX	1664753	1664499
BMA_A1536	type III secretion system protein BsaW	1665469	1664789
BMA_A1537	type III secretion system protein BsaV	1666442	1665459
BMA_A1540	type III secretion system protein BsaS	1669435	1668125
BMA_A1541	type III secretion system protein BsaR	1669839	1669432
BMA_A1542	type III secretion system protein BsaQ	1671923	1669851
BMA_A1543	type III secretion system protein BsaP	1673080	1671959
BMA_A1544	type III secretion system protein BsaO	1674897	1673077
BMA_A1545	type III secretion system transcriptional regulator BsaN	1675726	1674968
BMA_A1547	type III secretion system protein BsaM	1676120	1677406
BMA_A1548	type III secretion system protein BsaL	1677403	1677672
BMA_A1550	type III secretion system BasJ	1678034	1678984
BMA_A1551	type III secretion apparatus protein OrgA/MxiK	1678981	1679568
BMA_A1552	type III secretion apparatus protein, HrpE/YscL family	1679537	1680331
BMA_A1602	type III secretion outer membrane pore, YscC/HrcC family	1740918	1739119
BMA_A1613	type II/III secretion system family protein	1754093	1752303
BMA_A1625	type III secretion inner membrane protein, authentic frameshift	1763834	1762891
BMA_A1627	type III secretion inner membrane protein SctS	1765159	1764896
BMA_A1628	type III secretion inner membrane protein SctR	1765850	1765200
BMA_A1629	type III secretion inner membrane protein SctQ	1767135	1765837
BMA_A1630	type III secretion inner membrane protein SctV	1769843	1767771
BMA_A1631	type III secretion protein, YscU/HrpY family	1770922	1769840
BMA_A1632	type III secretion protein, HrpB1/HrpK family	1771158	1771733
BMA_A1633	type III secretion protein HrpB2	1771747	1772163
BMA_A1634	type III secretion inner membrane protein SctJ, authentic frameshift	1772166	1773012
BMA_A1635	type III secretion protein HrpB4	1773009	1773686
BMA_A1636	type III secretion inner membrane protein SctL	1773671	1774387
BMA_A1637	type III secretion apparatus H+-transporting two-sector ATPase	1774423	1775712

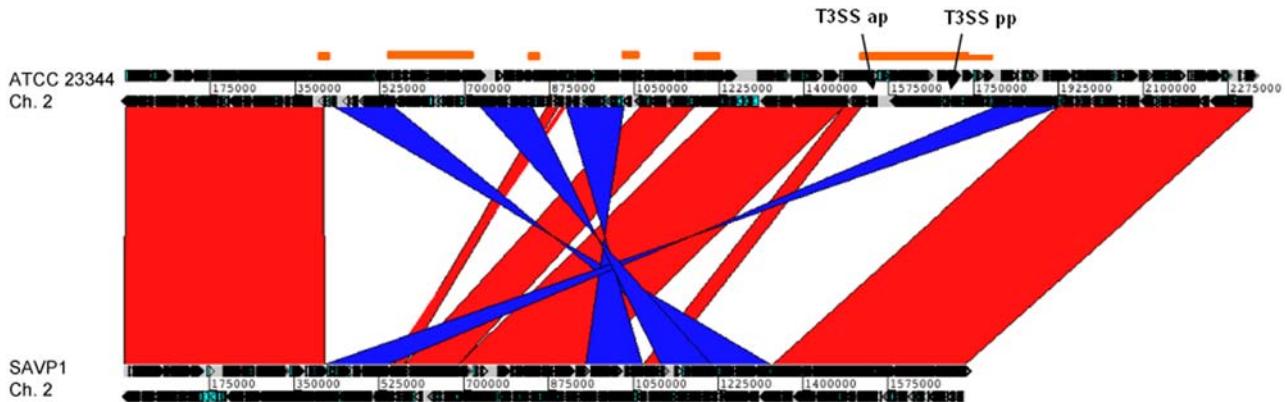
All genes are located on chromosome II and appear to be present in two basically contiguous segments: BMA\_A1520-BMA\_A1552 and BMA\_A1625-BMA\_A1637. Genes with Bsa are part of the animal pathogen-like T3SS. A complete list of the genes present in ATCC 23344 but missing from SAVP1 is given in Supplementary Table S1. doi:10.1371/journal.pone.0002058.t001

segments at insertion sequences (IS), particularly IS407A elements (Fig. 1 and 2). SAVP1 has lost approximately 610 kb of DNA due to IS mediated deletions (see Supplemental Table 1 for genes not present in SAVP1). Most of the lost genes are on chromosome 2, a replicon that encodes more accessory functions than chromosome 1. Notable in comparing presence of genes on SAVP1 but not on ATCC 23344 include multiple fimbriae/pili genes and several hypothetical genes (See Figure 2). The rearrangements detailed here demonstrate the plasticity of the *B. mallei* genome and suggest that IS mediated deletions may have a profound effect on the relative virulence of *B. mallei* strains.

Thus, based on our current knowledge, the strain is unlikely to undergo a natural restoration to a virulent state by passage in an animal host, in contrast to the situation if there were only reparable replication errors or point mutations in a single or a few genes.

## Discussion

The general public remains at risk to the pathogenic effects of *B. mallei* until we can develop reliable therapeutics, vaccines or other protective measures, and rapid diagnostics. We would have an opportunity to develop these countermeasures more quickly if we could identify attenuated strains of *B. mallei* and employ the strains in selected experiments. Unlike anthrax, where we have effective diagnostics, approved therapy, and vaccines, none of these resources are available for *B. mallei*. This means that those most capable of advancing research in this field are at high occupational risk, without effective contingency measures, should an infection be suspected or actually occur. In addition to occupational risk, two other factors that impede research are the requirement to use BSL-3 facilities and the Select Agent regulations[23,24]. Fortunately, there are provisions within the Select Agent Program that



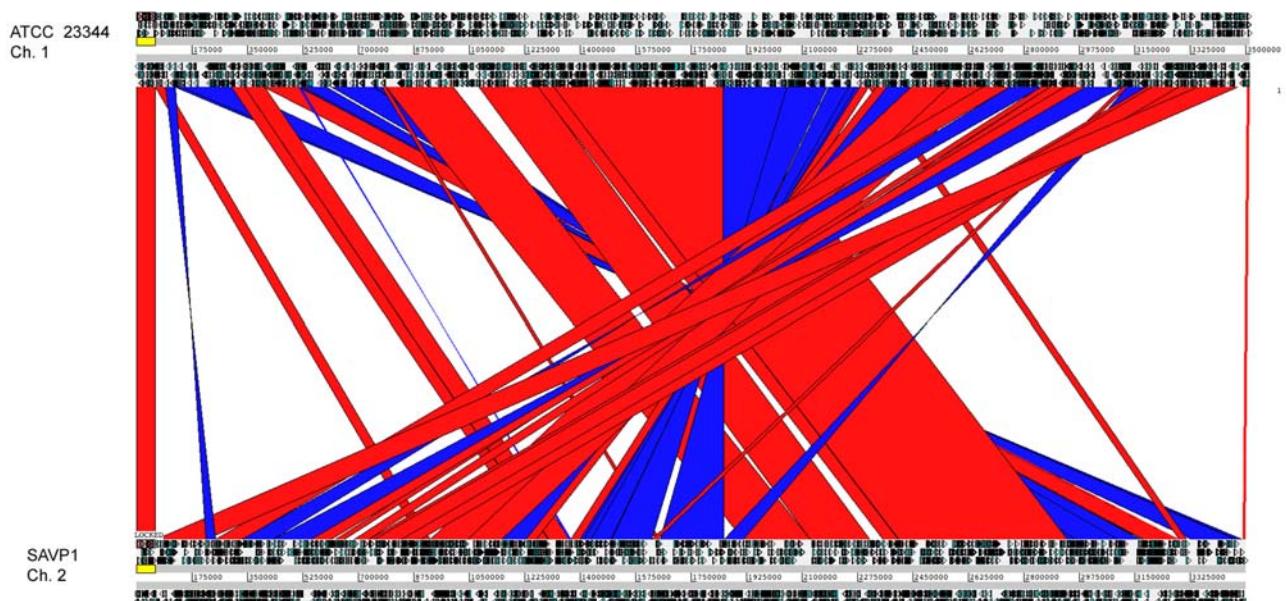
**Figure 1. Comparison of the small chromosomes of *B. mallei* ATCC 23344 (top) and *B. mallei* SAVP1 (bottom) using ACT and the whole genome alignment MUMmer.** Regions of similarity, rearrangements, and deletions are readily apparent between these two strains. The red and blue bands represent the forward and reverse matches, respectively. The orange horizontal bars represent the regions of ATCC 23344 that are missing in SAVP1. The locations of the animal pathogen-like T3SS (T3SS ap) and the plant pathogen-like (T3SS pp) gene clusters that are present in ATCC 23344, but absent in SAVP1, are indicated by arrows. The ATCC 23344 small chromosome (CP0000525) is 1.73 Mb.

doi:10.1371/journal.pone.0002058.g001

enable a particular microbial strain to removed or excluded[25]. Certain strains of microbes have already been excluded from the list and not subject to the requirements of 42 CFR Part 73 and 9 CFR Part 121 if used in basic or applied research, as positive controls, for diagnostic assay development, proficiency testing, or for the development of vaccines and therapeutics. Examples now excluded are certain strains of *Yersinia pestis*, *Bacillus anthracis* strains devoid of both plasmids pX01 and pX02, *Bacillus anthracis* strains devoid of the plasmid pX02 (e.g., *Bacillus anthracis* Sterne, pX01+pX02-), *Brucella abortus* Strain 19 and strain RB51 (vaccine strains), *Coxiella burnetii*, *Francisella tularensis* subspecies *novicida*, and *Francisella tularensis* subspecies *holartica* LVS (live vaccine strain)[25].

However, the regulations go back into effect if there is any reintroduction of factor(s) associated with virulence or other manipulations that restore the virulence or diminish the attenuation. Therefore, as might be done with excluded strains above, we believe it is prudent to monitor for any possible reversions with the SAVP1 during *in vivo* experiments. The same caveat applies to restoration experimentation, if it could be done, to provide more definitive evidence of that absence of T3SS as necessary and sufficient for attenuation.

We believe it is preferable to pair the strain with clinical cases in a natural-host infection whether that occurs naturally or by experiment. The closer the strain isolate is to the case the better,



**Figure 2. Comparison of the large chromosomes of *B. mallei* ATCC 23344 (top) and *B. mallei* SAVP1 (bottom) using ACT and the whole genome alignment MUMmer.** Regions of similarity, rearrangements, and deletions are readily apparent between these two strains. The red and blue bands represent the forward and reverse matches, respectively. The ATCC 23344 large chromosome (CP000010) is 3.51 Mb and the SAVP1 large chromosome (CP0000526) is 3.49 Mb.

doi:10.1371/journal.pone.0002058.g002

rather than an isolate from multiple *in vitro* passages in cultures. Misleading data might result by randomly selecting an isolate from a strain collection without the associated clinical information. In the current example we happened to identify the loss of animal T3SS genes in a strain associated with avirulent phenotypic behavior in its natural host. Because the T3SS is so essential to export of virulent factors, and this was unexpectedly found by WGS of SAVP1, there is increased confidence that experiments with this isolate may be accomplished with a higher degree of safety than with the virulent strain. Loss of some of the other 582 ORFs could be additive to this effect. Presence of other genes which are on SAVP1 and not on ATCC 23344 could also contribute to attenuated behavior. Though this strain would no longer be our choice for comparative genomic studies to detect subtle genetic differences that may account for virulence, its ability to evoke antibody responses confers it with great potential for other types of biodefense research as well as potential vaccines in equids and humans. Other avirulent strains may also prove to be candidates for Select Agent regulation exclusion[16,17].

In summary, our investigation into virulence factors using WGS on clinically-associated strains of *B. mallei* led to an unexpected finding. This finding may serve to eventually have at least one strain, SAVP1, removed from select agent constraints. Its research utility could be assessed. We believe there are comparable situations with other biothreat agents. We hope that the example of our finding with SAVP1 will engender discussion among public health and regulatory agencies, academia, and the private sector that will favorably impact our biodefense research efforts.

## Materials and Methods

### Sequencing

The genomes of *B. mallei* were sequenced and assembled by random shotgun method as described[22].

### Coding Sequence (CDS) Prediction and Gene Identification

Open reading frames (ORFs) likely to encode proteins (CDSs) were identified by using GLIMMER. Identified CDSs were annotated by manual curation of the outputs of a variety of similarity searches. Searches of the predicted coding regions were performed with BLASTP, as described[26]. The protein–protein matches were aligned with blast\_extend\_repraze, a modified Smith-Waterman[27] algorithm that maximally extends regions of similarity across frameshifts. Gene identification is facilitated by searching against a database of nonredundant bacterial proteins (nraa) developed at The Institute for Genomic Research (TIGR) and curated from the public archives GenBank, Genpept, Protein

Information Resource, and SwissProt. Searches matching entries in nraa have the corresponding role, gene common name, percent identity and similarity of match, pairwise sequence alignment, and taxonomy associated with the match assigned to the predicted coding region and stored in the database. CDSs were also analyzed with two sets of hidden Markov models constructed for a number of conserved protein families from PFAM and TIGRFAM. Regions of the genome without CDSs and CDSs without a database match were reevaluated by using BLASTX as the initial search, and CDSs were extrapolated from regions of alignment. Finally, each putatively identified gene was assigned to one of 113 role categories.

### ATCC 23344 - SAVP1 Comparative Analysis

All CDSs from ATCC 23344 were aligned against the whole genome sequence of SAVP1 with the Program to Assemble Spliced Alignments (PASA) [28]. PASA first summons BLAT[29] to align the CDSs to the genome and then validates each alignment by requiring a minimum 95% sequence identity over at least 90% of the gene length. Alignments failing BLAT validation are then realigned using sim4[30] and revalidated using the same criteria. All ATCC 23344 CDSs that could not be aligned were thus assumed to be absent from SAVP1. Similar analyses were applied to reverse strain comparisons.

### Animal Studies

The experimental infection of equids was performed under the auspices and regulations of the United States Department of Agriculture on Plum Island.

### Supporting Information

**Table S1** Table of Genes present in ATCC23344 but absent from SAVP1

Found at: doi:10.1371/journal.pone.0002058.s001 (0.08 MB PDF)

### Acknowledgments

We express our gratitude to Wolfgang F. Fricke for assistance with design of the figures.

### Author Contributions

Conceived and designed the experiments: CF SS BL LS JD. Performed the experiments: LS WN CR. Analyzed the data: CF SS JR DD BL LS JD WN CR. Contributed reagents/materials/analysis tools: DD LS WN. Wrote the paper: CF SS DD BL LS JD WN CR. Other: Figure construction: JR.

## References

- Wilkinson L (1981) Glanders: medicine and veterinary medicine in common pursuit of a contagious disease. Medical History 25: 363–384.
- Whitlock GC, Estes DM, Torres AG (2007) Glanders: off to the races with *Burkholderia mallei*. FEMS Microbiol Lett 277: 115–122.
- CDC (2000) Glanders (*Burkholderia mallei*) http://www.cdc.gov/ncidod/dbmd/diseasinfo/glanders\_g.htm.
- CDC et al (2000) Laboratory-acquired human glanders—Maryland, May 2000. MMWR Morb Mortal Wkly Rep 49: 532–535.
- Fraser CM (1991) The Merck veterinary manual—a handbook of diagnosis, therapy, and disease prevention and control for the veterinarian. Rahway, NJ: Merck & Co.
- Mandell GL, Douglas RG, Bennett JE (1995) Mandell, Douglas and Bennett's principles and practice of infectious diseases. New York: Churchill Livingstone.
- Geissler E, Moon JEvC, Stockholm International Peace Research Institute (1999) Biological and toxin weapons research, development and use from the Middle Ages to 1945. Oxford: Oxford University Press.
- Wheelis M (1998) First shots fired in biological warfare. Nature 395: 213.
- Congressional Research Service (CRS), Shea DA, Gortron F (2004) Small-scale Terrorist Attacks Using Chemical and Biological Agents (updated)(also at www.fas.org/irp/crs/RL32391-062304.pdf). RL32391.
- Department of Health and Human Services ASIPaR (2007) http://www.hhs.gov/aspr/barda/documents/phemce\_implplan\_041607final.pdf.
- Koenig RL (2006) The Fourth Horseman: One Man's Secret Campaign to Fight the Great War in America. New York: Public Affairs.
- Franz DR, Jahrling PB, Friedlander AM, McClain DJ, Hoover DL, et al. (1997) Clinical recognition and management of patients exposed to biological warfare agents. JAMA 278: 399–411.
- Kortepeter M, Christopher GW, Cieslak T, Culpepper R, Darling R, et al. (2001) USAMRIID's Medical management of biological casualties handbook. Fort Detrick: U.S. Army Medical Research Institute of Infectious Diseases.
- Fritz DL, Vogel P, Brown DR, DeShazer D, Waag DM (2000) Mouse model of sublethal and lethal intraperitoneal glanders (*Burkholderia mallei*). Vet Pathol 37: 626–636.

15. Ulrich RL, DeShazer D (2004) Type III secretion: a virulence factor delivery system essential for the pathogenicity of *Burkholderia mallei*. *Infect Immun* 72: 1150–1154.
16. DeShazer D, Waag DM, Fritz DL, Woods DE (2001) Identification of a *Burkholderia mallei* polysaccharide gene cluster by subtractive hybridization and demonstration that the encoded capsule is an essential virulence determinant. *Microb Pathog* 30: 253–269.
17. Schell MA, Ulrich RL, Ribot WJ, Brueggemann EE, Hines HB, et al. (2007) Type VI secretion is a major virulence determinant in *Burkholderia mallei*. *Mol Microbiol* 64: 1466–1485.
18. Yabuuchi E, Kosako Y, Oyaizu H, Yano I, Hotta H, et al. (1992) Proposal of *Burkholderia* gen. nov. and transfer of seven species of the genus *Pseudomonas* homology group II to the new genus, with the type species *Burkholderia cepacia* (Palleroni and Holmes 1981) comb. nov. *Microbiology & Immunology* 36: 1251–1275.
19. Srinivasan A, Kraus CN, DeShazer D, Becker PM, Dick JD, et al. (2001) Glanders in a military research microbiologist. *N Engl J Med* 345: 256–258.
20. Fritz DL, Vogel P, Brown DR, Waag DM (1999) The hamster model of intraperitoneal *Burkholderia mallei* (glanders). *Veterinary Pathology* 36: 276–291.
21. Lopez J, Copps J, Wilhelmsen C, Moore R, Kubay J, et al. (2003) Characterization of experimental equine glanders. *Microbes Infect* 5: 1125–1131.
22. Nierman WC, DeShazer D, Kim HS, Tettelin H, Nelson KE, et al. (2004) Structural flexibility in the *Burkholderia mallei* genome. *Proc Natl Acad Sci U S A* 101: 14246–14251.
23. Centers for Disease Control and Prevention (2005) Select Agent Program. <http://www.cdc.gov/od/sap/>.
24. US Department of Agriculture A (2005) Agricultural Bioterrorism Protection Act of 2002; Possession, Use, and Transfer of Biological Agents and Toxins. Federal Register 70: 13242–13292.
25. Centers for Disease Control and Prevention (2006) Select Agent Program Notification of Exclusion. <http://www.cdc.gov/od/sap/sap/exclusion.htm>.
26. Fleischmann RD, Adams MD, White O, Clayton RA, Kirkness EF, et al. (1995) Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* 269: 496–512.
27. Waterman MS (1988) Computer analysis of nucleic acid sequences. *Methods Enzymol* 164: 765–793.
28. Haas BJ, Delcher AL, Mount SM, Wortman JR, Smith RK Jr, et al. (2003) Improving the *Arabidopsis* genome annotation using maximal transcript alignment assemblies. *Nucleic Acids Res* 31: 5654–5666.
29. Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, et al. (2002) The human genome browser at UCSC. *Genome Res* 12: 996–1006.
30. Florea L, Hartzell G, Zhang Z, Rubin GM, Miller W (1998) A computer program for aligning a cDNA sequence with a genomic DNA sequence. *Genome Res* 8: 967–974.